DOCUMENT RESUME

ED 167 528 SP 013 897

AUTHOR Malitz, David; And Others
TITLE Selecting Subjects to Be Trained to Use Observational Systems: A Replication and Extension. Technical Report No. 5064.
INSTITUTION Texas Univ., Austin. Research and Development Center for Teacher Education.
PUB DATE 78
NOTE 59p.

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Classroom Observation Techniques; *Evaluation Criteria; Individual Characteristics; Interaction Process Analysis; Lesson Observation Criteria; *Performance Factors; Predictive Ability (Testing); *Project Training Methods; Reliability; Self Concept; *Training Techniques

ABSTRACT

The object of this research was to establish a reliable test battery to determine probable future success for those persons who wish to be trained as expert classroom observers. The focus of this investigation was to predict coding performance. This report presents the findings of a replication of an earlier study that indicated that coder performance could be predicted from measures of training performance and coder attitudes toward classrooms. These findings were partially confirmed. Included in this report are results of tests conducted to assess accuracy in coding and reliability of observation. Also included are samples of a self-concept test, and two Likert-type questionnaires on attitudes about classroom behavior and beliefs and attitudes about teaching and the teaching process. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Evaluation of Teaching

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED167524

SELECTING SUBJECTS TO BE TRAINED TO USE OBSERVATIONAL SYSTEMS: A REPLICATION AND EXTENSION

David Malitz, Richard O. Clements
and Linda K. Zimmerer

Technical Report No. 5064

November 30, 1978

The Research & Development Center
for Teacher Education

The University of Texas at Austin

UTR&D

SP 013 241

Selecting Subjects to be Trained to Use Observational Systems:

A Replication and Extension

Classroom observation systems have become a popular means of quantifying teacher and student behavior in recent years. These systems are widely used both as research instruments and as means of providing formal or informal feedback to inservice teachers and trainees. The degree of interest in the use of classroom observation techniques is evidenced by the sizable number of articles devoted to this topic which are currently documented in the ERIC system (approximately 200). Yet despite this burgeoning interest, there is little in print which can serve as a guide to the novice researcher or practitioner who wishes to implement a classroom observational system. Details seem especially lacking in how to train coders to use these systems. Even manuals or technical reports which are available for some systems are rather reticent in this respect, leaving the novice in the position of either hiring a consultant or devising his or her own training procedures. Furthermore, even where one is able to obtain "expert" advice on various aspects of training, it is often the case that this advice is based upon word-of-mouth or informal observation rather than upon empirical evidence.

One of these aspects of training which is potentially crucial is, "Who should be trained?" Is it best to train people who are knowledgeable and experienced in education, or are naive subjects sufficient? Or, is it possible that naive subjects are more objective than knowledgeable ones since they have fewer preconceived notions about what should or should not occur in classrooms? Perhaps certain skills (such as clerical ability) or personality attributes are relevant to coding accuracy or objectivity.

In the past few years, the Evaluation of Teaching program (EOT) at the Research and Development Center for Teacher Education at The University of Texas at Austin has been directly involved in training over 100 subjects in the use of various classroom observation systems. In that time, two facts have become evident. First, some people learn to code more accurately than others. And, second, no data exist to aid the trainer in predicting, either at or before the beginning of training, which subjects are most likely to become competent coders. Because training can be time consuming and expensive, it would be useful to know how to predict coding performance. If predictors of coding performance could be isolated, accurate coders could be selected. In this way training time and effort could be spent on those most likely to perform well and the overall quality of the data obtained from the study's coders could be improved over that expected from an unscreened group of coders. For this reason, the EOT program undertook a preliminary study (Malitz, 1978) to identify predictors of coding performance. The results of that study (termed "Phase I" in this report) were promising and served as the basis for the current study (termed "Phase II") reported here. The procedures and results of Phase I are reviewed below (for a detailed description, see Malitz and Clements, 1978; Marston, Kugle, and Zimmerer, 1978). A detailed description of the procedures and results of the Phase II study are presented below.

Phase I Study

Two potential predictors of coding performance were identified in discussions among EOT staff members who had served as trainers and/or coders. The first of these was coding performance after several hours of training. Since it is typical for trainees to code samples of behavior early in training in order to test their understanding of the observation system employed, data from reliability checks, as these tests are called, was easily obtainable.

The first reliability check, for example, can be viewed as a sample of on-the-job performance and therefore carries potential for predicting later performance.

The second predictor variable identified by the EOT staff was coders' attitudes towards teachers, students, and classrooms. Because no observation system can capture completely all of the complexities of classroom behavior, ambiguities frequently exist within these systems. Moreover, it is impossible for training to prepare coders for every situation they will encounter while observing classroom behavior. For this reason coders frequently must rely upon their own judgment in coding ambiguous behaviors. It was hypothesized that in these ambiguous situations coders' attitudes towards teachers and students might affect their coding performance.

The opportunity to test these notions about coder performance came when 30 subjects were hired to participate in another study (Marston, Kugle, and Zimmerer, 1978). These subjects received seven hours of training in the use of the Flanders System of Interaction Analysis (FSIA) (Flanders, 1971). Twenty-six subjects completed training and went on to code 11 half-hour videotapes of fifth-grade social studies classes in which teacher used traditional instructional methods.

Coder performance was operationally defined as each coder's average reliability in terms of agreement with the other coders across all of the 11 tapes. To compute this average, a matrix of reliabilities for each of the 11 tapes was produced in which the coefficient of reliability of each coder with every other coder was determined (Scott, 1955). Once these matrices were generated, it was possible to compute each coder's mean reliability with every other coder on each of the tapes. Each coder's means were then

averaged across the tapes, resulting in a final measure of coder performance, termed the coder's "average reliability." To have obtained a high average reliability across all of the tapes, a coder must have demonstrated high reliability with a relatively large number of the other coders on the tapes. Coders who had a low average reliability were those who had difficulty attaining high reliability with other coders.

The first predictor of coding performance was based on data obtained during the first of two reliability checks conducted over the course of training. This reliability check was conducted at the end of the first training day after approximately three hours of training. Each coder's average reliability on this training tape, computed in the manner described above, served as a measure of his or her early training performance. The second predictor was a measure of coder attitudes towards teachers, students, and classrooms. This measure consisted of coders' scores on a shortened version of the Minnesota Teacher Attitude Inventory (MTAI) (Cook, Leeds, and Callis, 1951) administered before the beginning of training. This shortened version (MTAI-S) consisted of 35 items taken from the original 150-item MTAI. These 35 items were chosen by EOT staff members as representative of the content areas sampled by the full MTAI.

With coders' average reliabilities across the 11 tapes as criterion, correlations were computed between the criterion and each of the two predictors. Significant correlations ($p < .05$) were obtained for both the MTAI-S scores ($r = .45$) and training reliabilities ($r = .48$). Further analysis demonstrated that use of the MTAI-S as a screening device in this sample of subjects would have resulted in significantly more reliable coding than was obtained from the entire unscreened group. This was considered very encouraging because of the ease and economy with which the MTAI-S can be administered.

Phase II Study

The basic purpose of the Phase II study was to replicate the above findings using a second classroom observation system, the OScAR 5V (Medley, 1976; Medley & Mitzel, 1958). In addition, this study was designed to remedy several of the shortcomings of the earlier investigation: (1) various attitude, personality, and ability measures were administered in addition to the MTAI-S in an attempt to identify more clearly what coder characteristics are and are not related to coder performance; and (2) the coding performance data were obtained from a videotape coded after training, but before the coders' attitudes were experimentally manipulated for another study. The Phase I criterion data were obtained in the midst of another experiment. Although statistical analysis showed no interaction between the Phase I findings and the tandem experiment's manipulations, the present study was designed to produce data untainted by any such manipulations. A third purpose of the Phase II study was to present some evidence relating to the validity of "average reliability" as a measure of coder performance. The Phase I study relied heavily upon this index both as a predictor and a criterion. As a measure of coder performance, average reliability suffers from the fact that it is not measured against an objective standard of "good" or "correct" coding. Instead, it is a relative measure in which the best coder is defined as the one who achieves the highest reliability with all other coders. From a trainer's point of view, this is a very useful index of performance. A person who scores high on this index, if randomly paired with another coder from the same training group, would likely achieve high interrater agreement. Despite this practical appeal, there is no assurance, in the absence of supporting data, that "average reliability" is in any sense a meaningful index of coder performance. A more common

and more generally accepted index of a coder's performance is that coder's degree of agreement with an expert coder on a sample of classroom behavior. In order to lend some credence to the "average reliability" measure, the Phase II study was designed so that some "expert codings" were available. This made it possible on certain videotapes to compute for each coder both his or her average reliability and the degree of agreement with an expert. This allowed the two indices, average reliability and agreement with expert, to be compared with one another.

Method

Coders

Coders were recruited through an ad in the university student newspaper. They were paid $3 per hour for their training, plus a $20 bonus for being on time for all the experimental sessions. None of the coders had previous teaching or classroom observation experience, nor were they education majors. They came from a broad range of backgrounds. The experiment was conducted during the regular university spring term. Thirty-five subjects were hired to participate in the study. However, only 25 coders participated in the training program, and five of them failed to complete all the sessions. Two additional coders dropped after completing the training due to scheduling problems. Thus, only 18 of the original 35 subjects (51%) completed training and participated in the study.

Coding Instrument

The Observation Schedule and Record Form (OScAR 5V) (Medley & Mitzel, 1958) was selected for this experiment. It is a system for classifying teacher verbal behaviors observed in the classroom. The coding instrument used for this experiment was a slightly modified version of the OScAR 5V (See Appendix A for sample form used in this study). The instrument is

teacher-focused; coders are instructed to categorize the teacher verbalizations on the basis of the student's perception of the teacher's behavior. Coding on the OScAR 5V is event-based, an event being defined as an utterance or sequence of utterances which express a single idea or thought (Medley, 1976). Two different types of events are recognized: statements and interchanges. Both statements and interchanges may be classified as substantive (focus on lesson content) or non-substantative (focus on subjects other than lesson content). In turn, interchanges may be subdivided into entry or exit utterances (statements are always classified as entries). An entry is an utterance which is initiating and, therefore, not a response to a previous event. Entries may be emitted either by the teacher or a pupil. An example of an entry behavior is "teacher asks question." In contrast, exits are always teacher responses to preceding pupil entries. Exits include such behavior as positive or negative affective responses.

Coder Training

Two well-trained OScAR coders who had previously been employed by EOT were hired as consultants to train the newly recruited coders for this experiment. The trainers were both graduate students in educational psychology who originally had been taught to use the OScAR by the system's author. Following this initial training, they coded more than 40 hours of videotaped classroom interaction for another EOT study (Borich, Malitz, & Kugle, 1978). During that time they achieved consistently high reliability with one another.

The training sessions were conducted over a five-day period for three hours each. Pretests and posttests were administered daily to assess subjects' understanding of the material presented each day. The system was described through the use of category descriptions, tables and flow charts. Coders

practiced with the use of typed scripts, audiotapes, and videotapes of classroom behavior.

Training Reliability

During the course of training, four reliability checks were conducted to assess coders' progress in mastering the system. During these checks, conducted at the beginning and end of the fourth and fifth days of training, coders were asked to code portions of videotapes from the same series as the experimental tapes described below. In all cases, reliability measures were based on the intraclass correlation coefficient (Winer, 1971). The results of these reliability checks are reported in Table 1. Two reliability indices are reported in that table. The first index is the group's average intraclass reliability across all coders. To compute this index, intraclass coefficients were calculated between all possible pairings of coders, and the resulting coefficients were averaged across the pairings. This index is termed "average within-group reliability" in Table 1. The second index, "average reliability with expert", was calculated for three of the four reliability checks. On each of these, the trainers carefully coded the test videotapes. For each of these tapes, reliabilities were computed between each coder's protocol and the experts' codes, and the resulting intraclass coefficients were averaged. Note that no average reliability with the experts is reported for the second reliability check. This is because time constraints prevented the experts from coding this tape.

The data presented in Table 1 suggest that the training procedures were

successful in producing reliable coders. There was a general trend towards higher reliability on the fifth (final) day of training as compared with the fourth. Moreover, the level of reliability by the end of training, for both indices, was quite high, especially considering the complexity of the 67-category OScAR system.

Videotapes

The videotapes used for this experiment were originally produced for the Individualized Teaching for Effective Coping study (Haak, Lawrence, & Peck, 1971). These tapes were made in traditionally structured public school classrooms at the fifth- and seventh-grade levels for the purpose of providing feedback to teachers. The tapes were made in existing classes, and the teachers were encouraged to be as natural as possible. The tapes were all 25-30 minutes in length. A thirty-minute tape of math instruction was selected as the criterion tape. Five-, fifteen-, and thirty-minute segments of tapes of social studies teachers were selected for training. The videotapes were selected on the basis of technical quality as well as subject matter.

Questionnaires

A number of questionnaires were administered to subjects during the week of training. Some were completed during the training sessions while others were taken home and returned later in the week. These questionnaires measured a variety of personality, attitude, demographic and miscellaneous variables which staff members speculated might be related to coder behavior. Following is a description of each of the instruments:

(1) The Adjective Checklist (ACL) (Gough & Heilbrun, 1965) consists of a list of 300 adjectives commonly used to describe attributes of a person. Subjects were instructed to read each adjective quickly and to place a check beside each one they considered self-descriptive. The ACL manual provides 24 scoring keys which can be applied to ACL protocols. These keys produce scores on a variety of variables including such traits as self-confidence, dominance, aggression, etc. The reader is referred to Appendix B for a list of the ACL scales and to Gough and Heilbrun (1965) for a full description of them. Protocols were scored according to the scale keys and the resulting errors were converted to standard scores using tables provided in the ACL Manual. These standard scores were used to investigate relationships between personality and coder behavior. A copy of the ACL is provided in Appendix C of this report.

(2) The Questionnaire for Coders (QFC) was developed by EOT staff members and is included in Appendix D. The first page of the QFC provides a brief introduction to the EOT research project, solicits demographic information, and questions subjects about previous experience with coding systems and teaching.

The remainder of the QFC is labeled "Classroom Attitude Survey" (CAS) and consists of 56 statements about classrooms,teachers, and students to which statements subjects indicate their degree of agreement on a five-point scale. The first 35 statements are items taken from the Minnesota Teacher Attitude Inventory (MTAI) (Cook, Leeds, & Callis, 1951). These 35 items were chosen as representative of the content areas surveyed by the MTAI. These items were used in an earlier study (Malitz, 1978) and were scored as on the full MTAI. The resulting scores were found to be predictive of later coding performance. Scores on the 35 MTAI items were used in the current study in an attempt to replicate these earlier findings.

The remaining 21 items on the QFC were classroom attitude items written

by the EOT staff. Analysis of data collected in an earlier study (Marston, Kugle, & Zimmerer, 1978) showed these items to be related in a complex manner to coder behavior. It was hypothesized that these items and the MTAI items might share content in common and that factor scores might produce relationships with coder behavior more easily interpretable than analyses of individual items. For this reason, an unpublished study was conducted by the EOT staff in which the entire set of 56 items was administered to a group of approximately 300 psychology undergraduates. Their ratings were factor analyzed and four factors emerged. These four factors were named Teacher Responsiveness, Teacher Control, Student Ability, and Valence of Educational Experience. Protocols were scored according to this factor structure and these scores were used to further investigate the relationships between coder attitudes and behaviors. Note that in Appendix D the items which loaded onto each of these four factors are indicated.

(3) A third classroom attitude questionnaire was completed by the coders. This survey was developed by Wehling and Charters (1969) and can be found in Appendix E of this report. It contains 86 statements to which subjects indicate their degree of agreement on a seven point scale. Wehling and Charters administered this questionnaire to eight samples of teachers and student teachers and found eight factors: (1) Subject-matter emphasis, (2) personal adjustment ideology, (3) student autonomy vs. teacher direction, (4) emotional disengagement, (5) consideration of student viewpoint, (6) classroom order, (7) student challenge, and (8) integrative learning. Subjects' responses were scored according to the factor structure described by Wehling and Charters (1969).

(4) It was hypothesized that clerical ability might be a predictor of coder performance since various clerical skills are necessary to use

observation systems properly. For this reason, the Minnesota Clerical Test (MCT) (Andrew & Paterson, 1946) was administered to coders. This is a speed and accuracy test consisting of two parts, of which only the first part, "number checking," was used for this study. The number checking test consists of 200 pairs of numbers ranging from three to 11 digits. Subjects are given eight minutes in which to successively compare the two numbers in each pair and place a check mark on a line if they are identical or to leave the line blank if the two numbers differ. A copy of this test is included in Appendix F. Subjects' protocols were scored for the number of items completed and the number of items completed correctly. Andrew and Paterson (1946) report that these two scores are highly correlated (r = .96) since most items completed are completed correctly (approximately 95%).

Procedure

At the end of the final day of training, subjects were instructed to return for another session several days later. At the beginning of this final session, the 18 subjects present viewed and coded a half-hour videotape from the series described above. This tape will be referred to as the "criterion tape." After coding this tape, coders completed a six-item post-coding questionnaire devised by EOT staff members. This questionnaire surveyed coders' subjective impressions of the teacher they had viewed. A copy of this instrument is included in Appendix G. Following this, subjects received an experimental attitude manipulation which served as the basis for another study (Coder Bias Tech. Report, 1979). However, all of the data reported here were collected before the final experimental manipulation took place.

Results

For the 18 subjects who completed training and coded the criterion tape, average reliabilities were computed on the criterion tape and on the four

reliability checks in the manner described above. In addition, reliabilities with expert codings were computed for reliability checks one, three, and four. Finally, all questionnaire data were scored for these 18 subjects. Below are the results of analyses involving these variables. These results will be presented in four sections. The first will present data demonstrating the validity of the average reliability index. The second section will present the results of analyses intended to replicate the results of the Phase I study. The third section will present analyses involving additional variables examined in this study but not represented in the Phase I study. The final section analyzes differences which emerged between subjects who completed the study and those who dropped out.

Validity of the Average Reliability Index

The reader's attention is directed to Table 2 , which contains Pearson product-moment correlations demonstrating the validity of the average reliability index. Three of the correlations in this table are relevant to this issue. Expert codings were available for reliability checks one, three, and four (variables A, B, and C in Table 2). Correlations were computed between each coder's average reliability on each of these occasions (variables D, F, and G) and his or her reliability with the expert's codes. It can be seen that the correlations for these three occasions were .87, .88, and .82, all significant ($p < .001$). Expert codings of tapes are traditionally considered the standard against which a trainee's codes can be judged for accuracy. However, these high correlations suggest that a trainee's performance will be judged in a very similar manner on the basis of the average degree of agreement achieved with his or her peers.

Predictive Validity of Phase I Variables

It will be recalled that in the Phase I study, both performance during

training and attitude scores on the MTAI-S were found to be predictive of later coding performance. In the present study, seven measures of training performance were available for analysis (three reliability-with-expert measures (variables A through C) and the average reliability measures (variables D through G). For both sets of scores, it appeared that reliability checks 1 and 3 were more predictive of criterion reliability than checks 2 and 4. This pattern is rather interesting, for checks 1 and 3 were conducted at the beginning of training days 4 and 5 before the day's training began, while checks 2 and 4 occurred at the end of each training day. It should be recalled that the criterion tape was coded several days after the last day of training. If this pattern of correlations was not a chance occurrance, it would indicate that performance after the completion of training was best predicted by performance at the beginning of a training session, rather than after several hours of drill and instruction. One might attribute this pattern to fatigue on the part of the coders at the end of the training session. If the coders were fatigued, then performance measures taken at that time might have lacked validity. However, reference to the training reliabilities reported in Table 1 reveals no evidence of fatigue; the reliabilities at the end of the training sessions were in fact quite high.

Instead, this pattern of results might best be explained by a recall effect. Criterion performance was dependent upon how much instruction the coder retains after training has ended. Training performance measured at the beginning of a training session was dependent upon recall of instruction during the previous days of instruction, while performance measured at the end of a session may have been more dependent upon immediate recall. Viewed in this way, the performance situation at the beginning of a training session

was more similar to the criterion situation than was the situation at the end of a training session. Therefore, it is not unreasonable to expect these pre-session measures to be more predictive of later performance than the post-session measures.

It should be noted, however, that the correlations among the average training reliabilities did not support this line of reasoning. For if the pre-session situations were similar to one another, one would expect the two pre-session average reliabilities (variables D and F in Table 2) to have been highly correlated with one another. Inspection of the correlations among the four training average reliability measures (variables D through G) showed this not to be the case. In fact, of all of the correlations among these four measures, the highest (and the only significant) correlation was between a pre-session and a post-session measure (variables E and F, $r = .66$).

Nevertheless, the basic finding of the Phase I study, that criterion performance can be predicted by training performance, was generally confirmed in the present study. However, it appeared that while measures of training performance at certain points in time were highly predictive of later performance, measures taken at other points in time predicted less efficiently. The data available from the present study did not allow any firm conclusions as to why this might be so.

The second finding of the Phase I study was that scores on the MTAI-S were predictive of criterion performance. The correlation between the MTAI-S (variable H) and the criterion reliability (variable I) in Table 2 was low and non-significant ($r = .13$, $p > .05$)., clarly not confirming the previous finding. Interesting, however, the MTAI-S scores were correlated significantly with performance on the fourth reliability check as measured by both reliability-with-expert ($r = .59$, $p < .05$) and average reliability

(r = .58, p < .05). However, correlations with performance on the other three checks were all non-significant.

Thus, of the two findings of the Phase I study, one was replicated in the present study. Training performance was found to be predictive of criterion performance. However, coder attitudes, as measured by the MTAI-S, were not found to be predictive of this performance.

Predictive Validity of Additional Variables

A number of predictor variables, not measured in the Phase I study, were correlated with criterion performance. This section presents the results of these correlations.

The Adjective Checklist (ACL) (Gough and Heilbrun, 1965) can be scored to produce 24 different scales (See Appendix C). Scores on these 24 scales were computed for each of the 18 subjects who coded the criterion tape, and correlations were computed between each of these sets of scores and criterion reliability. Of these 24 correlations, only two (8.3%) were significant (p < .05). The first of these two significant correlations was between criterion performance and the number of unfavorable adjectives checked (r = -.40). The direction of this correlation indicated that subjects who performed well checked fewer unfavorable adjectives. The dimension of "favorableness-unfavorableness", as used by Gough and Heilbrun, appears to be similar to the more familiar dimension of social desirability. According to Gough and Heilbrun, subjects who check few of these words may be characterized as "more placid, more obliging, more mannerly, more tactful, and probably less intelligent" (p. 8).

The second significant correlation with the ACL was with the "abasement" scale, defined as the tendency "to express feelings of inferiority through self-criticism, guilt, or social impotence" (Gough and Heilbrun, p. 11). Effective coders tended to score low on this scale ($r = -.44$) and may be characterized as "optimistic, poised, productive, and decisive" (Gough and Heilbrun, p. 11).

These two scales on the ACL were the only additional variables which correlated significantly with coders' criterion performance. Correlations with the eight factors of Wehling and Charters' (1969) attitude survey, with the four factors of the attitude survey developed by EOT staff, and with two scores derived from the Minnesota Clerical Test (Andrew and Paterson, 1946) were all non-significant ($p > .05$). In addition, age, sex, and educational level of the coders were found to be unrelated to performance.

In summary, 41 variables were correlated with criterion performance, and only the two ACL scale scores produced significant correlations. Since only 4.9% of the variables were found to be significant at the 5% level, there is little reason to believe that a non-chance association existed between this group of variables and the criterion.

Comparison of Completers and Drop-Outs

As has been shown above, the three measures of classroom attitudes (the MTAI-S, the Wehling and Charters questionnaire, and the EOT questionnaire) failed to correlate with criterion performance. This finding was at odds with the finding of the Phase I study in which the MTAI-S scores were found to be related to performance. In considering why this difference might have occurred, it should be noted that a high proportion of the Phase II subjects who were hired (49%) either failed to show up for training or failed to complete

training. This high drop-out rate was not present in the Phase I study, where 26 of 28 hired subjects (93%) completed the study. It was therefore possible that the drop-out process was not a random one and that the lack of relationship between the attitude measures and performance might have been due to some systematic bias present in the subjects who completed the study.

All 35 Phase II subjects who were hired completed several questionnaires before they were hired, allowing comparison of completers and drop-outs on several variables. These variables were MTAI-S scores, scores on the four factors of the EOT attitude questionnaire (EOTQ-I through EOTQ-IV), age, educational level, and sex. All of these variables (with the exception of sex) are continuous, and a series of *t*-tests were performed to test for differences between completers and drop-outs on each of these variables. These tests were performed using the SPSS T-TEST program, and the results of these analyses are presented in Table 3. For each variable in Table 3, the means and standard deviations for the two groups are presented, as are the associated *t*-values, degrees of freedom, and two-tailed probabilities. For one variable (EOTQ-I), an analysis of variance suggested that the variances for the two populations differed from one another. Since the assumption of equal population variances is crucial to the *t*-test, an alternative test was performed. This test is described in the SPSS manual (Nie, Hull, Jenkins, Steinbrenner, and Bent, 1975, pp. 269-270). This test produces a statistic which is approximately distributed as the *t*-distribution, and it is based on separate variance estimates for the two groups, rather than a pooled estimate. It also involves the use of a formula which produces an estimate of appropriate degrees of freedom which can be rounded to the nearest whole number.

These analyses showed that completers and drop-outs differed

significantly on three of the five attitude variables measured: MTAI-S, EOTQ-I, and EOTQ-III. The means for the two groups on the MTAI-S indicated that the drop-outs scored lower than the completers. In the Phase I study, high MTAI-S scores were associated with better coding performance. Therefore, it appeared that subjects who dropped out tended to have low MTAI-S scores, and that the remaining subjects who completed the study had high scores. This might have restricted the range of the MTAI-S scores among completers, thereby attenuating the correlation possible between the MTAI-S and criterion performance. The means on EOTQ-I and EOTQ-III indicated that drop-outs tended to score lower on these two scales as well. These two factors have been named "teacher responsiveness" and "student ability." Low scores on these two factors suggested endorsement of classrooms in which teachers did not tailor their teaching to student needs and indicated a lack of faith that students were capable of thinking, reasoning, and making decisions on their own. These low scores were consistent with low MTAI-S scores which suggested endorsement of autocratic and controlled, rather than democratic and permissive classrooms.

Returning to Table 3, it should be noted that the completers and drop-outs did not differ in age or educational level. A chi-square analysis was performed to test for an association between sex and drop-out. The results of this analysis are presented in Table 4 where it can be seen that the chi-square value was non-significant.

In summary, significant differences in attitude scores were found between subjects who completed and those who dropped out of the study. Completers were characterized by more liberal classroom attitudes which, in the Phase I study, were associated with better coding performance. These findings suggest that

the lack of relationship found between attitudes and performance among completers may have been due to a restricted range of attitudes among completing subjects.

The focus of this investigation has been to predict coding performance. However, from a practical standpoint it would be equally useful to be able to identify subjects who will not complete training. It was therefore of interest to assess the predictive efficiency of the variables available in identifying those subjects who dropped out.

A discriminant analysis was performed using the SPSS DISCRIMINANT program in order to quantify this predictive efficiency. A binary vector was constructed which contained either a "1" or a "0" for each of the 35 hired subjects: a "1" if the subject completed the study, and a "0" if he or she did not. This binary vector served as the discriminant variable. The number of subjects was considered too small to effectively use discriminant analysis as an inferential tool to assess the relative importance of the variables at hand. Instead, it was used as a descriptive tool to assess the predictive strength of the three variables for which *t*-tests had shown the two groups of subjects (completers and drop-outs) to differ significantly (MTAI-S, EOTQ-I, and EOTQ-III). Therefore, these three variables were entered as discriminating variables in the equation with the binary vector as the discriminant variable.

The discriminant function which emerged as the solution to this equation weighted the three predictors approximately equally and was found to have an eigenvalue of .36 ($p = .014$). The canonical correlation associated with this function was .53. When applied to this sample, the function correctly classified 74.3% of the subjects. Thus, on the basis of scores on these three attitude variables, one could predict correctly for about three-fourths

of the sample whether a subject would complete or drop out of the study. Thus, for this sample at least, these variables appeared to have some practical utility in identifying drop-outs.

Discussion

The primary purpose of this investigation was to replicate the basic findings of the Phase I study: that coder performance could be predicted from measures of training performance and coder attitudes toward classrooms. These findings were partially confirmed. Of seven measures of training performance, four were significantly correlated with criterion performance. The pattern of correlations obtained suggested that criterion performance may be more accurately predicted by performance on reliability checks made at the beginning, rather than the end, of particular training sessions. However, the data were not entirely consistent with this conclusion, and the data available prevented any definitive conclusions about when training performance should be measured.

Thirteen measures of coder attitudes were correlated with criterion performance, and none of these correlations reached significant levels. One of these measures (the MTAI-S) was the same measure used in the Phase I study. Thus, the Phase I conclusion that coder attitudes were related to coder performance was not confirmed in this study. However, it was noted that nearly half of the coders hired for this study failed to complete it. Furthermore, it was demonstrated that subjects who completed the study could be differentiated from those who dropped out on the basis of their scores on several attitude measures (including the MTAI-S). In fact, these measures demonstrated moderate predictive validity in classifying completers and drop-outs. It was hypothesized that the abnormally high attitude scores among completers restricted the range of these measures,

thereby attenuating the correlations obtained between attitude scores and criterion performance.

The high attitude scores among completers suggested that these subjects endorsed a democratic permissive approach to teaching and dealing with children. In the Phase I study, similar attitudes were associated with high coding performance. These findings suggested that subjects who profess this set of attitudes (let us call them "permissive attitudes" for lack of a better phrase) make better candidates for training than subjects with "less permissive" attitudes. The more permissive subjects are more likely to complete training, and, once training is completed, they tend to code more reliably. If one were interested solely in prediction and in selecting training candidates, these findings would be satisfactory as they stand. However, it is not unreasonable to ask why subjects' scores on these questionnaires were related to coding performance.

It was hoped that the additional variables collected in this study would help answer this question by showing more specifically what coder characteristics are and are not related to coder performance. However, as a group these variables (personality measures, demographic data, and clerical aptitude scores) failed to achieve more significant correlations with the criterion than one would expect to find by chance. Only two of these variables, two scale scores on the Adjective Checklist, correlated with the criterion, and it is difficult to understand how the characteristics purportedly measured by these scales could have been related to coding performance. Thus, there was little data available in this study with which to explain the relationship between coder performance and coder attitudes.

Several explanations for this finding can be offered, although these explanations are rather speculative. One possibility is that, when faced with ambiguous

coding situations, coders' attitudes help them decide which category most appropriately captures the behavior observed. If this were true, then one would expect that coders with similar attitudes, be they permissive or restrictive, would tend to code in a similar manner. This similarity in coding would be reflected in high reliability with coders who hold similar beliefs. However, in the Phase I study the data tended to disconfirm this notion. In that study, coders were divided into two groups of approximately equal size: those who scored high on the MTAI-S (HI-M), and those who scored low (LO-M). Average reliabilities on each of the eleven criterion tapes were computed separately for each of these groups. This analysis showed that even though coders in the LO-M group presumably held similar attitudes, they consistently achieved low reliability with one another. Thus, similar attitudes did not bring about similar coding; coders who scored low on the MTAI-S tended to disagree on coding with one another as well as with coders who scored high on the MTAI-S. The Phase II study provided no additional data on this point since no association was found between attitudes and performance among subjects who completed the study. However, the Phase I findings did cast some doubt upon the hypothesis that attitudes directly affect coding.

Another possible explanation is that coders who scored in the more permissive direction on the attitude questionnaires were somehow more motivated as coders. Perhaps because they were more motivated, these coders attended the training sessions more diligently and concentrated more on learning the coding system and using it correctly. This hypothesis is consistent with the findings of both the Phase I and Phase II studies, for it explains why coders with the more liberal attitudes were more likely to complete the Phase II study and why they coded more reliably in Phase I. Although this motivation hypothesis

explains the data, it raises another question: why would scores on classroom attitude instruments reflect motivation to code? It is possible that this motivation stemmed from an interest in children. As one examines the items on these attitude instruments (see appropriate appendices), one notes that many of the items seem to assess the degree to which the respondent likes children, has faith in their abilities, and believes that adults should conform their behavior to children's needs. Perhaps a person who endorses these attitudes is one who is more interested in watching videotapes of classrooms and in learning a system of coding classroom behaviors.

Although this hypothesis is tenable, it is based upon many assumptions about both the attitude instruments and the effect of interest patterns on motivations. Basically, the difficulty in interpreting the Phase I and Phase II findings stemmed from ambiguity about what was measured by the attitude instruments. If they were truly measuring classroom attitudes, then one must provide a plausible model for the influence of attitudes on coding. Unfortunately, the findings that LO-M coders tended to disagree with one another as well as with HI-M coders made it difficult to construct such a model. One is then forced to assume that the instruments measured some other construct (such as interest or motivation). Plausible models can be constructed based on these constructs which are consistent with the data. But the assumptions about what the instruments measured are at odds with the questionnaires' design and must be made without any supporting evidence. Thus, while one may speculate about the reasons for the findings in these studies, it does not appear possible to provide an unequivocal explanation based upon the data available.

The above comments have all been directed toward the main purpose of this study (the replication of the Phase I study). However, it will be recalled that a secondary purpose of this study was to provide evidence regarding the

validity of the average reliability index as a measure of coding performance. In a philosophical sense, such validity can never truly be established, for it would require a perfectly accurate coding of classroom behavior against which one could compare coders' protocols. Unfortunately, such perfect coding does not exist since the observation systems themselves only imperfectly capture the observed behavior. No matter how careful and knowledgeable one is, there will always be situations which cannot be coded unambiguously. From a practical standpoint, however, expert codings are generally accepted as a reasonable substitute for perfect coding. Experts in a particular system generally agree most of the time about how behavior should be coded. An expert's protocol therefore is used as the standard against which trainees are compared. If a trainee's protocol agrees with the expert's, his or her performance is considered good. Low reliability with the expert is interpreted as poor performance. If the average reliability index rank orders subjects in a manner similar to the reliability-with-expert index, then it can be considered a valid index of performance. In fact, the data presented in this study strongly suggested that in this sense the average reliability index is a valid measure of performance. The two indices were so highly correlated that they were virtually interchangeable in the way in which they rank ordered coders' performance. It should be noted, however, that with a small group of coders average reliability may lose some of its validity since a few deviant coders who agreed with one another could achieve high average reliability even though their codes might not be correct. Nevertheless, average reliability is a useful and easily computed index and can probably be used safely in other studies, at least if it is computed among more than a handful of coders.

References

Andrew, D. M., & Patterson, D. G. Minnesota Clerical Test manual. New York: The Psychological Corp., 1946.

Borich, G. D., Malitz, D., and Kugle, C. L. Convergent and discriminant validity of five classroom observational systems: Testing the model. Journal of Educational Psychology, 1978, 70, 119-128.

Clements, R. O., Malitz, D., & Zimmerer, L. K. The effects of situational bias on coding performance. Paper to be presented at the annual meeting of the American Educational Research Association, San Francisco, California, 1979.

Cook, W. W., Leeds, C. H., & Callis, R. Minnesota Teacher Attitude Inventory: Manual. New York: The Psychological Corporation, 1951.

Flanders, N. A. Analyzing teacher behavior. Reading, Mass.: Addison-Wesley, 1971.

Haak, R. A., Lawrence, W. A., & Peck, R. F. ITEC: An experiment in school consultation. Publication No. 1082. Austin, Texas: The University of Texas Research and Development Center for Teacher Education, 1971.

Gough, H. L., & Heilbrun, A. B. The Adjective Checklist manual. Palo Alto, California: Consulting Psychologists Press, 1965.

Malitz, D. Identification of predictors of coding performance: Selecting subjects to be trained to use observation systems. Paper presented at the annual meeting of the American Educational Research Association, Toronto, Canada, 1978.

Malitz, D., & Clements, R. O. Identification of predictors of coding performance: Selecting subjects to be trained to use observation systems. (Technical Report) Austin, Texas: University of Texas Research and Development Center for Teacher Education, 1978 .

Marston, P. T., Kugle, C. L., & Zimmerer, L. K. Coder drift: A reliability problem for teacher observation. (Technical Report) Austin, Texas: The University of Texas Research and Development Center for Teacher Education, 1978.

Medley, D. M. Observation Schedule and Record (OScAR 5V) coder's manual. Draft of May, 1976. (Mimeo)

Medley, D. M., & Mitzel, H. E. A technique for measuring classroom behavior. Journal of Educational Psychology, 1958, 49, 86-92.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. Statistical package for the social sciences (2nd ed.). New York: McGraw-Hill, 1975.

Scott, W. A. Reliability of content analysis: The case of nominal scale coding. The Public Opinion Quarterly, 1955, 19, 321-325.

Wehling, L. J., & Charters, W. L. Dimensions of teacher beliefs about the teaching process. American Educational Research Journal, 1969, 6, 7-30.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1971.

Table 1

Results of Reliability Checks Conducted During Training

Check	Day of Training	Number of Coders Present	Average Within-Group Reliability[a]	Average Reliability With Expert[a]
1	beginning of 4th	19	.62	.67
2	end of 4th	20	.64	--[b]
3	beginning of 5th	20	.79	.85
4	end of 5th	20	.72	.78

[a] see text for explanation of these reliability indices

[b] expert coding not available for this tape

Table 2

Intercorrelations of Training Reliabilities, MTAI-S Scores, and Criterion Reliabilities

	A	B	C	D	E	F	G	H	I
A. Reliability with expert Check 1	---	.42	.20	.87*	.42	.15	.03	-.08	.56*
B. Reliability with expert Check 3		---	.16	.57*	.62*	.88*	.20	-.19	.69*
C. Reliability with expert Check 4			---	.28	.12	-.05	.82*	.59*	.30
D. Average reliability Check 1				---	.38	.39	.20	-.06	.73*
E. Average Reliability Check 2					---	.66*	.06	-.17	.31
F. Average reliability Check 3						---	.20	-.28	.61*
G. Average reliability Check 4							---	.58*	.40
H. MTAI-S								---	.13
I. Average reliability Criterion tape									---

* $p < .05$, one-tailed test.

Note: For variables A and D, n = 16. For variable E, n = 17. For all other variables, n = 18.

Table 3

Analysis of Differences Between Subjects Who

Completed and Dropped out of the Study

Variable	Completers (n = 18) Mean	S.D.	Drop-outs (n = 17) Mean	S.D.	T value	df	Two-tailed probability
MTAI-S	7.83	5.16	1.82	6.42	3.06	33	.004
EOTQ-I	37.39	1.91	34.59	4.54	2.35[a]	21.25[a]	.028
EOTQ-II	20.89	2.72	20.94	2.66	-.06	33	.955
EOTQ-III	21.00	1.94	18.76	2.66	2.85	33	.007
EOTQ-IV	18.00	2.95	17.71	3.08	.29	33	.775
AGE	21.94	2.75	22.29	4.07	-.30	33	.767
EDUC.	3.72	1.78	3.29	1.65	.74	33	.466

[a]T-value, df, and probability for EOTQ-I based on an approximation of the t-distribution used for populations with unequal variances. See text for explanation.

Table 4

Chi-square analysis of sex for subjects who completed and dropped out of study

	Completers	Drop-outs	
Males	7 (9.25)	11 (8.74)	18 (17.99)
Females	11 (8.74)	6 (8.27)	17 (17.01)
	18 (17.99)	17 (17.01)	35 (35.00)

$\chi^2 = 1.41$, df = 1, p = .234

Note: Theoretical frequencies in parentheses. Chi-square value corrected for continuity.

Appendix A

OScAR Handout

The basic unit of observation on which a record is based is called the utterance. Any verbalization which expresses or communicates an idea is called an utterance. An utterance consists of a single word or a long sentence or series of words. So long as it comes across to the observer as expressing a single idea or thought, it may be recorded as a single utterance regardless of length.

The task of the coder may be described simply as the classification and recording of two types of verbalizations: entries and exits.

An entry is an utterance which is initiating, one which is not a response to a preceding utterance; it may be emitted either by the teacher or by a pupil. Pupil utterances elicited by and directly responsive to teacher utterances are not regarded as initiating and are not recorded.

An exit, then, is a teacher utterance responsive to a pupil utterance.

TABLE 1

Behaviors Recorded in the First Word (ENTRIES)

	Symbol	Name	Brief Description
1.	PNS	Pupil Utterance Non-Substantive	Pupil makes a statement or asks a question not related to substance
2.	PQU	Pupil Question Substantive	Pupil asks for substantive information
3.	PST	Pupil Statement Substantive	Pupil offers substantive information
4.	PRS	Pupil Response	Pupil responds directly to another pupil or indirectly to teacher
5.	PBST	Problem Structuring Statement	Teacher raises a substantive question or sets a problem (without indicating who is to answer or solve it)
6.	CVG	Convergent Question	Teacher asks pupil a question which calls for one right answer
7.	EL	Elaborating Question	Teacher directs question to pupil whose answer depends on the preceding one
8.	DVG	Divergent Question	Teacher asks pupil a question to which more than one answer may be acceptable or correct

TABLE 2

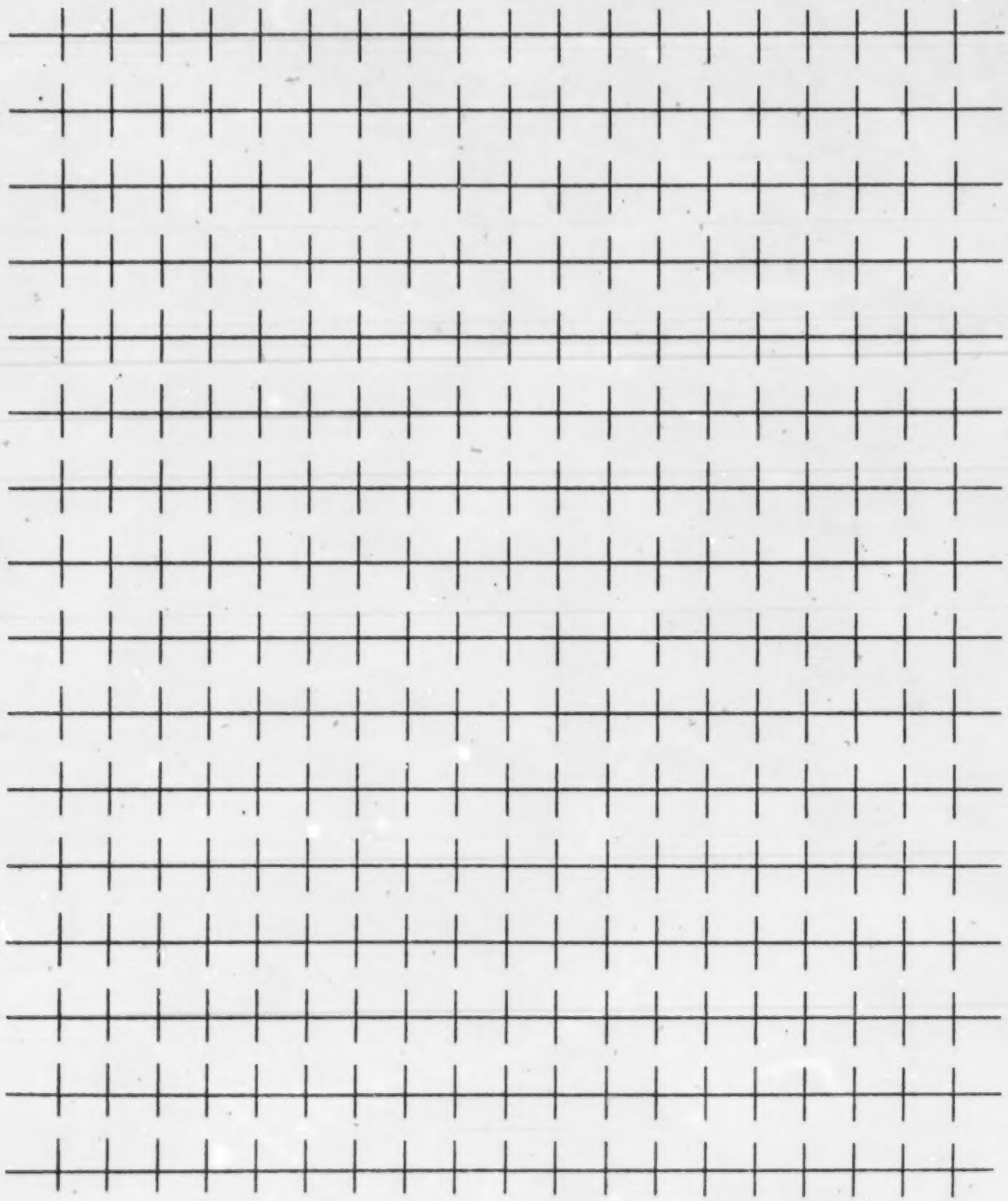
Behaviors Recorded in the Second Word (EXITS)

	Symbol	Name	Brief Description
1.	NOEV	No Evaluation	Teacher does not reply to pupil utterance
2.	CNSUP	Considering-Supporting	Teacher utterance with positive affect
3.	INFAP	Informing-Approving	Teacher gives information or positive feedback
4.	DSCAC	Describing-Accepting	Teacher accepts pupil response or makes statement not otherwise classifiable
5.	DIREJ	Directing-Rejecting	Teacher commands pupil to do something or gives negative feedback
6.	RBCRT	Rebuking-Criticizing	Teacher utterance with negative affect
7.	DST	Desisting	Teacher commands pupil to stop doing something or refuses permission
8.	PRNS	Procedural, Neutral-Non-Substantive Question	Teacher asks question not otherwise classifiable; teacher neither refuses nor gives permission
9.	PR+	Procedural, Positive	Teacher offers pupil choice of action or gives permission

OSCAR CODING SHEET

Name of Coder_____ Date_____

Tape Number_____ Time_____



Appendix B

Adjective Checklist Scales

1. Total number of adjectives checked
2. Defensiveness
3. Number of favorable adjectives checked
4. Number of unfavorable adjectives checked
5. Self-confidence
6. Self-control
7. Lability
8. Personal adjustment
9. Achievement
10. Dominance
11. Endurance
12. Order
13. Intraception
14. Nurturance
15. Affiliation
16. Heterosexuality
17. Exhibition
18. Autonomy
19. Aggression
20. Change
21. Succorance
22. Abasement
23. Deference
24. Counseling readiness

Appendix C
BACS-1-Spring 1978

_____
Name

DIRECTIONS FOR USING BACS-1 ANSWER SHEET

This answer sheet contains a list of 300 adjectives. Please read them quickly and place a check beside each one you would consider to be self-descriptive. (Example: [✓]). Try to be frank and check the adjectives which describe you as you really are, not as you would like to be. Be sure and continue through adjective No. 300.

(Copyright 1952 by Harrison G. Gough, Ph.D., University of California, Berkeley. Published by Consulting Psychologists Press, 577 College Ave., Palo Alto, California.)

ID_____

1.	[ ]	absent-mirded	51.	[ ]	cowardly	101.	[ ]	gloomy
2.	[ ]	active	52.	[ ]	cruel	102.	[ ]	good-looking
3.	[ ]	adaptable	53.	[ ]	curious	103.	[ ]	good-natured
4.	[ ]	adventurous	54.	[ ]	cynical	104.	[ ]	greedy
5.	[ ]	affected	55.	[ ]	daring	105.	[ ]	handsome
6.	[ ]	affectionate	56.	[ ]	deceitful	106.	[ ]	hard-headed
7.	[ ]	aggressive	57.	[ ]	defensive	107.	[ ]	hard-hearted
8.	[ ]	alert	58.	[ ]	deliberate	108.	[ ]	hasty
9.	[ ]	aloof	59.	[ ]	demanding	109.	[ ]	headstrong
10.	[ ]	ambitious	60.	[ ]	dependable	110.	[ ]	healthy
11.	[ ]	anxious	61.	[ ]	dependent	111.	[ ]	helpful
12.	[ ]	apathetic	62.	[ ]	despondent	112.	[ ]	high-strung
13.	[ ]	appreciative	63.	[ ]	determined	113.	[ ]	honest
14.	[ ]	argumentative	64.	[ ]	dignified	114.	[ ]	hostile
15.	[ ]	arrogant	65.	[ ]	discreet	115.	[ ]	humorous
16.	[ ]	artistic	66.	[ ]	disorderly	116.	[ ]	hurried
17.	[ ]	assertive	67.	[ ]	dissatisfied	117.	[ ]	idealistic
18.	[ ]	attractive	68.	[ ]	distractible	118.	[ ]	imaginative
19.	[ ]	autocratic	69.	[ ]	distrustful	119.	[ ]	immature
20.	[ ]	awkward	70.	[ ]	dominant	120.	[ ]	impatient
21.	[ ]	bitter	71.	[ ]	dreamy	121.	[ ]	impulsive
22.	[ ]	blustery	72.	[ ]	dull	122.	[ ]	independent
23.	[ ]	boastful	73.	[ ]	easy-going	123.	[ ]	indifferent
24.	[ ]	bossy	74.	[ ]	effeminate	124.	[ ]	individualistic
25.	[ ]	calm	75.	[ ]	efficient	125.	[ ]	industrious
26.	[ ]	capable	76.	[ ]	egotistical	126.	[ ]	infantile
27.	[ ]	careless	77.	[ ]	emotional	127.	[ ]	informal
28.	[ ]	cautious	78.	[ ]	energetic	128,	[ ]	ingenious
29.	[ ]	changeable	79.	[ ]	enterprising	129.	[ ]	inhibited
30.	[ ]	charming	80.	[ ]	enthusiastic	130.	[ ]	initiative
31.	[ ]	cheerful	81.	[ ]	evasive	131.	[ ]	insightful
32.	[ ]	civilized	82.	[ ]	excitable	132.	[ ]	intelligent
33.	[ ]	clear-thinking	83.	[ ]	fair-minded	133.	[ ]	interests narrow
34.	[ ]	clever	84.	[ ]	fault-finding	134.	[ ]	interests wide
35.	[ ]	coarse	85.	[ ]	fearful	135.	[ ]	intolerant
36.	[ ]	cold	86.	[ ]	feminine	136.	[ ]	inventive
37.	[ ]	commonplace	87.	[ ]	fickle	137.	[ ]	irresponsible
38.	[ ]	complaining	88.	[ ]	flirtatious	138.	[ ]	irritable
39.	[ ]	complicated	89.	[ ]	foolish	139.	[ ]	jolly
40.	[ ]	conceited	90.	[ ]	forceful	140.	[ ]	kind
41.	[ ]	confident	91.	[ ]	foresighted	141.	[ ]	lazy
42.	[ ]	confused	92.	[ ]	forgetful	142.	[ ]	leisurely
43.	[ ]	conscientious	93.	[ ]	forgiving	143.	[ ]	logical
44.	[ ]	conservative	94.	[ ]	formal	144.	[ ]	loud
45.	[ ]	considerate	95.	[ ]	frank	145.	[ ]	loyal
46.	[ ]	contented	96.	[ ]	friendly	146.	[ ]	mannerly
47.	[ ]	conventional	97.	[ ]	frivolous	147.	[ ]	masculine
48.	[ ]	cool	98.	[ ]	fussy	148.	[ ]	mature
49.	[ ]	cooperative	99.	[ ]	generous	149.	[ ]	meek
50.	[ ]	courageous	100.	[ ]	gentle	150.	[ ]	methodical

(continue to next page)

ID_____

151.	[ ]	mild	201.	[ ]	reliable	251.	[ ]	strong
152.	[ ]	mischievous	202.	[ ]	resentful	252.	[ ]	stubborn
153.	[ ]	moderate	203.	[ ]	reserved	253.	[ ]	submissive
154.	[ ]	modest	204.	[ ]	resourceful	254.	[ ]	suggestible
155.	[ ]	moody	205.	[ ]	responsible	255.	[ ]	sulky
156.	[ ]	nagging	206.	[ ]	restless	256.	[ ]	superstitious
157.	[ ]	natural	207.	[ ]	retiring	257.	[ ]	suspicious
158.	[ ]	nervous	208.	[ ]	rigid	258.	[ ]	sympathetic
159.	[ ]	noisy	209.	[ ]	robust	259.	[ ]	tactful
160.	[ ]	obliging	210.	[ ]	rude	260.	[ ]	tactless
161.	[ ]	obnoxious	211.	[ ]	sarcastic	261.	[ ]	talkative
162.	[ ]	opinionated	212.	[ ]	self-centered	262.	[ ]	temperamental
163.	[ ]	opportunistic	213.	[ ]	self-confident	263.	[ ]	tense
164.	[ ]	optimistic	214.	[ ]	self-controlled	264.	[ ]	thankless
165.	[ ]	organized	215.	[ ]	self-denying	265.	[ ]	thorough
166.	[ ]	original	216.	[ ]	self-pitying	266.	[ ]	thoughtful
167.	[ ]	outgoing	217.	[ ]	self-punishing	267.	[ ]	thrifty
168.	[ ]	outspoken	218.	[ ]	self-seeking	268.	[ ]	timid
169.	[ ]	painstaking	219.	[ ]	selfish	269.	[ ]	tolerant
170.	[ ]	patient	220.	[ ]	sensitive	270.	[ ]	touchy
171.	[ ]	peaceable	221.	[ ]	sentimental	271.	[ ]	tough
172.	[ ]	peculiar	222.	[ ]	serious	272.	[ ]	trusting
173.	[ ]	persevering	223.	[ ]	severe	273.	[ ]	unaffected
174.	[ ]	persistent	224.	[ ]	sexy	274.	[ ]	unambitious
175.	[ ]	pessimistic	225.	[ ]	shallow	275.	[ ]	unassuming
176.	[ ]	planful	226.	[ ]	sharp-witted	276.	[ ]	unconventional
177.	[ ]	pleasant	227.	[ ]	shiftless	277.	[ ]	undependable
178.	[ ]	pleasure-seeking	228.	[ ]	show-off	278.	[ ]	understanding
179.	[ ]	poised	229.	[ ]	shrewd	279.	[ ]	unemotional
180.	[ ]	polished	230.	[ ]	shy	280.	[ ]	unexcitable
181.	[ ]	practical	231.	[ ]	silent	281.	[ ]	unfriendly
182.	[ ]	praising	232.	[ ]	simple	282.	[ ]	uninhibited
183.	[ ]	precise	233.	[ ]	sincere	283.	[ ]	unintelligent
184.	[ ]	prejudiced	234.	[ ]	slipshod	284.	[ ]	unkind
185.	[ ]	preoccupied	235.	[ ]	slow	285.	[ ]	unrealistic
186.	[ ]	progressive	236.	[ ]	sly	286.	[ ]	unscrupulous
187.	[ ]	prudish	237.	[ ]	smug	287.	[ ]	unselfish
188.	[ ]	quarrelsome	238.	[ ]	snobbish	288.	[ ]	unstable
189.	[ ]	queer	239.	[ ]	sociable	289.	[ ]	vindictive
190.	[ ]	quick	240.	[ ]	soft-hearted	290.	[ ]	versatile
191.	[ ]	quiet	241.	[ ]	sophisticated	291.	[ ]	warm
192.	[ ]	quitting	242.	[ ]	spendthrift	292.	[ ]	wary
193.	[ ]	rational	243.	[ ]	spineless	293.	[ ]	weak
194.	[ ]	rattlebrained	244.	[ ]	spontaneous	294.	[ ]	whiny
195.	[ ]	realistic	245.	[ ]	spunky	295.	[ ]	wholesome
196.	[ ]	reasonable	246.	[ ]	stable	296.	[ ]	wise
197.	[ ]	rebellious	247.	[ ]	steady	297.	[ ]	withdrawn
198.	[ ]	reckless	248.	[ ]	stern	298.	[ ]	witty
199.	[ ]	reflective	249.	[ ]	stern	299.	[ ]	worrying
200.	[ ]	relaxed	250.	[ ]	stolid	300.	[ ]	zany

Thank you! You have completed the list.

Appendix D
EVALUATION OF TEACHING PROGRAM

QUESTIONNAIRE FOR CODERS

The Evaluation of Teaching program is conducting research in a number of areas with the goal of assessing effective teaching behavior. Your part in this effort will be to act as a trained observer of teacher behavior. By doing this, you can help us identify those teacher behaviors that are reliably coded with the system or systems you will be using.

As part of our research we would like to know what people of different backgrounds think about teachers and education. You can help us do this by filling out the attached Classroom Attitude Survey which consists of 40 statements relating to teachers, pupils, and the classroom. Your task is to rate on a scale of 1 to 5 how much you agree or disagree with each statement. If you strongly agree rate the statement "5", and if you strongly disagree rate it "1". The numbers between 1 and 5 represent less strong feelings. These statements are simply opinions and there is no "right" answer.

Before taking the attitude survey please fill out the questionnaire below. This information is necessary in determining whether people from different backgrounds have different attitudes towards the classroom situation. All of the information will be kept confidential.

Thank you for your help.

Name_____

Age_____ Education Level_____ Sex_____

Degrees_____ Subject_____Current Major_____

Do you know how to code with any other classroom observation systems? If the answer is yes, list the names of the systems.

( ) yes ( ) no

Have you ever worked as a teacher? If the answer is yes, give the grade level and number of years experience.

( ) yes ( ) no

Have you had any practice teaching experience?

( ) yes ( ) no

Do you intend to teach in the public schools?

( ) yes ( ) no

Classroom Attitude Survey

Circle the appropriate response for each item. "1" means that you strongly disagree and a "5" means that you strongly agree. The other numbers should represent gradations of these responses.

		Disagree				Agree
1.	The teacher should not be expected to manage a child if the latter's parents are unable to do so. (15)	1	2	3	4	5
2.	Discipline in the modern school is not as strict as it should be. (17)	1	2	3	4	5
3.	Standards of work should vary with the pupil. (24)	1	2	3	4	5
4.	Increased freedom in the classroom creates confusion. (33)	1	2	3	4	5
5.	There is too much emphasis on grading. (35)	1	2	3	4	5
6.	There is too great an emphasis upon "keeping order" in the classroom. (37)	1	2	3	4	5
7.	There are times when a teacher cannot be blamed for losing patience with a pupil. (48)	1	2	3	4	5
8.	Too many children nowadays are allowed to have their own way. (53)	1	2	3	4	5
9.	Most pupils are resourceful when left on their own. (62)	1	2	3	4	5
10.	Children should be allowed more freedom in their execution of learning activities. (71)	1	2	3	4	5
11.	Pupils must learn to respect teachers if for no other reason than that they are teachers. (72)	1	2	3	4	5
12.	Difficult disciplinary problems are seldom the fault of the teacher. (77)	1	2	3	4	5
13.	The whims and impulsive desires of children are usually worthy of attention. (78)	1	2	3	4	5
[D-] 14.	Children usually have a hard time following instructions. (79)	1	2	3	4	5

(Adapted from the Minnesota Teacher Attitude Inventory and reproduced by permission for research purposes. Copyright 1951 by the Psychological Corporation, New York, New York. All rights reserved.)

		Disagree				Agree
[C-] 15.	Children are unable to reason adequately. (83)	1	2	3	4	5
16.	Teachers who are liked best probably have a better understanding of their pupils. (89)	1	2	3	4	5
17.	Children should be given more freedom in the classroom than they usually get. (93)	1	2	3	4	5
18.	Pupils can be very boring at times. (98)	1	2	3	4	5
[C+] 19.	Most pupils are considerate of their teachers. (101)	1	2	3	4	5
20.	A teacher should never leave the class to its own management. (105)	1	2	3	4	5
21.	"Lack of application" is probably one of the most frequent causes for failure. (108)	1	2	3	4	5
22.	Grading is of value because of the competition element. (112)	1	2	3	4	5
[C-] 23.	Children usually will not think for themselves. (114)	1	2	3	4	5
24.	It isn't practical to base school work upon children's interest. (121)	1	2	3	4	5
25.	One should be able to get along with almost any child. (127)	1	2	3	4	5
[C-] 26.	Children are not mature enough to make their own decisions. (128)	1	2	3	4	5
[C+] 27.	Children will think for themselves if permitted. (130)	1	2	3	4	5
28.	Children should be given reasons for the restrictions placed upon them. (133)	1	2	3	4	5
[D-] 29.	Most pupils are not interested in learning. (134)	1	2	3	4	5
30.	It is usually the uninteresting and difficult subjects that will do the pupil the most good. (135)	1	2	3	4	5
[D+] 31.	Teachers should not expect pupils to like them. (141)	1	2	3	4	5

		Disagree				Agree
32.	Aggressive children require the most attention. (143)	1	2	3	4	5
33.	A pupil has the right to disagree openly with his teachers. (147)	1	2	3	4	5
34.	Most pupil misbehavior is done to annoy the teacher. (145)	1	2	3	4	5
[D-] 35.	One should not expect pupils to enjoy school. (149)	1	2	3	4	5
36.	Teachers should organize the class routine so that little time is lost in transition from one learning activity to another.	1	2	3	4	5
[B+] 37.	Effective teachers are consistent in their relationships with students.	1	2	3	4	5
38.	Most of my teachers have been competent.	1	2	3	4	5
[A+] 39.	A good teacher recognizes and responds to contributions of students.	1	2	3	4	5
[A+] 40.	Good teachers employ creative and imaginative approaches to teaching.	1	2	3	4	5
[D-] 41.	"Those who can, do, those who can't, teach," is a cliche that fits most elementary teachers.	1	2	3	4	5
[A+] 42.	Good teachers pace their instruction to suit students' concentration and interest span.	1	2	3	4	5
[B+] 43.	A good teacher maintains a neat personal appearance.	1	2	3	4	5
44.	Passing students to the next grade for social reasons is a good idea.	1	2	3	4	5
[B+] 45.	Teachers should design lessons and units with clear, logical and sequential content.	1	2	3	4	5
[A+] 46.	Teachers who utilize a variety of questioning strategies will have better results.	1	2	3	4	5
[A+] 47.	Teachers should communicate with students sympathetically, accurately and with understanding.	1	2	3	4	5

		Disagree				Agree
[B+] 48.	Effective teachers give careful attention to the physical condition and appearance of the classroom.	1	2	3	4	5
[A+] 49.	Showing initiative and imagination is an important teacher characteristic.	1	2	3	4	5
[A+] 50.	A teacher should involve all students in learning activities.	1	2	3	4	5
51.	You can't teach someone to teach--it's an inborn quality.	1	2	3	4	5
[A+] 52.	Practicing positive reinforcement techniques is an important aspect of teaching.	1	2	3	4	5
53.	An effective teacher prepares written lesson plans.	1	2	3	4	5
[B+] 54.	One cannot be a good teacher without having control over the classroom.	1	2	3	4	5
55.	A teacher must treat students impartially.	1	2	3	4	5
56.	An open classroom promotes learning.	1	2	3	4	5

Note: For the first 35 items, the numbers in parentheses indicate each item's original position on the MTAI. The letters to the left of item numbers indicate factor loadings (positive or negative) on the following four factors:

A. Teacher responsiveness
B. Teacher control
C. Student ability
D. Valence of educational experience.

Neither the factor loading indicators nor the numbers in parentheses appeared on the version of this questionnaire administered to subjects.

Appendix E

BACS-3 Spring 1978

The following 86 statements concern beliefs and attitudes about teaching and the teaching process. Please indicate your opinion about each of the following statements by circling a number from 1 to 7 to the right of it. If you strongly disagree, circle the low number. If you strongly agree with it, circle the high number. The numbers at the two ends of the scale represent the stronger feelings. Only mark a 1 or a 7 when your feelings are exceptionally strong. These two numbers should be used only rarely in responding to these statements. Please give an opinion about every statement even if you are not sure of its meaning.

Wehling, Leslie J., and Charter, W.W., Jr. Dimensions of Teacher Beliefs About the Teaching Process. American Educational Research Journal, Vol. 6, No. 1, January 1969.

		Disagree						Agree
1.	Teaching of specific skills and factual subject matter is the most important function of the school.	1	2	3	4	5	6	7
2.	The curriculum consists of subject matter to be learned and skills to be acquired.	1	2	3	4	5	6	7
3.	The backbone of the school curriculum is subject matter; activities are useful mainly to facilitate the learning of subject matter.	1	2	3	4	5	6	7
4.	Pupil failure is averted when mastery of subject matter is the prime requisite for promotion.	1	2	3	4	5	6	7
5.	The over-all plan of education suffers when teachers depart substantially from the subject outline.	1	2	3	4	5	6	7
6.	Grading pupils separately on achievement and citizenship assures that teachers will insist on mastery of subject matter as well as good behavior.	1	2	3	4	5	6	7
7.	Before pupils are encouraged to exercise independent thought they should be thoroughly grounded in the facts and knowledge about the subject.	1	2	3	4	5	6	7
8.	Learning is essentially a process of increasing one's store of information about various fields of knowledge.	1	2	3	4	5	6	7
9.	The structure of a field of knowledge is intrinsically interesting to pupils when it is clearly taught.	1	2	3	4	5	6	7
10.	In teaching, it is quite essential to cover the material in the course of study.	1	2	3	4	5	6	7
11.	The main reason for the curriculum guide is to provide the teacher with definite information regarding the material to be covered in the course.	1	2	3	4	5	6	7
12.	The essential function of junior high school courses lies in their preparing pupils for later courses.	1	2	3	4	5	6	7
13.	Pupils learn library skills more readily by using their own devices in searching for materials of special interest than by a series of exercises designed to teach the logical steps in library procedure.	1	2	3	4	5	6	7

		Disagree						Agree
14.	The teacher assures optimum learning conditions by giving top priority to the social-emotional needs of pupils.	1	2	3	4	5	6	7
15.	The development of social and emotional security for pupils is the most important function of the school.	1	2	3	4	5	6	7
16.	The individuality of pupils is sustained when teachers make allowances in their grade reports for the varying interest pupils have.	1	2	3	4	5	6	7
17.	Pupils gain a sense of belonging when the teacher encourages friendships among pupils in the room.	1	2	3	4	5	6	7
18.	Teachers increase their chances of directing the work into productive channels by having pupils participate in the planning.	1	2	3	4	5	6	7
19.	Group activity teaches children to think and plan together, independent of direct supervision by the teacher.	1	2	3	4	5	6	7
20.	The goals of education should be dictated by children's interests and needs as well as by the larger demands of society.	1	2	3	4	5	6	7
21.	The attitudes learned by a student are often the most important result of a lesson or unit.	1	2	3	4	5	6	7
22.	Small group work uses to best advantage the contrasting personalities, skills, and interests pupils have.	1	2	3	4	5	6	7
23.	Teachers who like pupils will usually encourage pupil initiation and participation in planning lessons.	1	2	3	4	5	6	7
24.	The effectiveness of the teacher depends entirely on the amount of personal interest he can invest in the progress of each pupil.	1	2	3	4	5	6	7
25.	Pupils master the essentials of a subject only when extensive plans are made for accommodating individual differences in pupils.	1	2	3	4	5	6	7
26.	A teacher can frequently "reach" a rebellious pupil by taking an intense personal interest in his welfare.	1	2	3	4	5	6	7
27.	Nothing stimulates a pupil to apply himself more diligently than a warm, personal interest in his progress shown by the teacher.	1	2	3	4	5	6	7

	Disagree						Agree
28. Teachers who do not like pupils will usually decide on and plan lessons alone rather than use pupil participation.	1	2	3	4	5	6	7
29. There is too great an emphasis on keeping order in the classroom.	1	2	3	4	5	6	7
30. Children should be given more freedom in the classroom than they usually get.	1	2	3	4	5	6	7
31. A properly motivated group of mature students might learn more in a semester's time if they were left entirely to their own resources than if they had a teacher to guide them.	1	2	3	4	5	6	7
32. Pupils frequently learn much more under their own initiative than they do under teacher direction.	1	2	3	4	5	6	7
33. Across-the-school routine imposes a consistency in classroom procedure which tends to restrict important avenues for learning.	1	2	3	4	5	6	7
34. Pupils are motivated to do better work when they feel free to move around the room while the class is in session.	1	2	3	4	5	6	7
35. Nothing captures students' interest in school work as quickly as allowing them to wrestle with problems of their own choosing.	1	2	3	4	5	6	7
36. When given a choice of activity, pupils generally select what is best for them.	1	2	3	4	5	6	7
37. Time to choose freely their own activity during the school day is a must for pupil morale.	1	2	3	4	5	6	7
38. Pupils learn best when permitted to set their own pace in doing the work.	1	2	3	4	5	6	7
39. A firm hand by the teacher promotes emotional security for pupils.	1	2	3	4	5	6	7
40. Pupils do their best work when they know exactly what to expect from day to day.	1	2	3	4	5	6	7
41. Pupils must be kept busy or they soon get into trouble.	1	2	3	4	5	6	7

	Disagree						Agree
42. Children need and should have more supervision and discipline than they usually get.	1	2	3	4	5	6	7
43. A well established classroom routine enhances the emotional stability of pupils.	1	2	3	4	5	6	7
44. Pupils must see clearly that it is the teacher, not they, who has charge of classroom learning.	1	2	3	4	5	6	7
45. Children learn the necessary skills of group participation only when they are exposed to sequences of activity requiring increasingly difficult skills from kindergarten through grade twelve.	1	2	3	4	5	6	7
46. The effective teacher has complete control of the learning situation at all times.	1	2	3	4	5	6	7
47. The teacher who organizes the material and presents it to pupils in a forceful way gets the best results.	1	2	3	4	5	6	7
48. Establishing the rules well in advance strengthens the teacher's hand in meeting the various problems that might arise.	1	2	3	4	5	6	7
49. A teacher's effectiveness rests upon his ability to maintain proper "professional distance" between the pupils and himself.	1	2	3	4	5	6	7
50. Pupils are induced to greater motivation when the teacher remains somewhat aloof from the interpersonal affairs of the class.	1	2	3	4	5	6	7
51. His effectiveness is seriously impaired when the teacher permits himself to become emotionally involved in the personal problems of pupils.	1	2	3	4	5	6	7
52. The effectiveness of teaching is enhanced when the teacher has the ability to see the world as each of his pupils sees it.	1	2	3	4	5	6	7
53. Students who misbehave or do not learn are generally children who need more love.	1	2	3	4	5	6	7
54. Children learn best in an atmosphere filled with love and emotional support.	1	2	3	4	5	6	7
55. The teacher's ability to see the world as each of his students sees it is an absolute must if he is to have any success at all in teaching.	1	2	3	4	5	6	7

		Disagree						Agree
56.	Good rapport with pupils is maintained by the teacher who always find time to help individuals with special problems.	1	2	3	4	5	6	7
57.	The use of sarcasm by the teacher can accomplish nothing but emotional harm for the pupil.	1	2	3	4	5	6	7
58.	Learning is enhanced when teachers praise generously the accomplishments of pupils.	1	2	3	4	5	6	7
59.	The pupil's impression of the teacher's personality greatly influences what he learns.	1	2	3	4	5	6	7
60.	Pupils learn to stay alert when they are expected to respond immediately to teacher demands.	1	2	3	4	5	6	7
61.	In the interest of good discipline pupils who repeatedly disrupt the class must be severely punished.	1	2	3	4	5	6	7
62.	Proper control of a class is amply demonstrated when pupils work quietly while the teacher is out of the room.	1	2	3	4	5	6	7
63.	Optimum learning takes place when the classroom setting is completely free of distractions.	1	2	3	4	5	6	7
64.	Under ideal conditions pupils would view each teacher as a "specialist" in the subject taught.	1	2	3	4	5	6	7
65.	Pupils learn efficiently the essentials of a subject when every member of the class moves simultaneously through carefully planned lesson sequences.	1	2	3	4	5	6	7
66.	The natural flow of events is enhanced by the teacher who manages to eliminate any disruptive pupil behavior.	1	2	3	4	5	6	7
67.	A good teacher will establish a routine and stick to it.	1	2	3	4	5	6	7
68.	The logical structure of subject matter is the most realistic guide to the organization of the work in the classroom.	1	2	3	4	5	6	7
69.	Pupils respect teachers who expect them to work hard in school.	1	2	3	4	5	6	7

		Disagree						Agree
70.	Lessons presented in the form of problems to be solved are the best means of motivating pupils.	1	2	3	4	5	6	7
71.	Pupils learn self direction by having opportunities to set their own goals for learning.	1	2	3	4	5	6	7
72.	Pupils respect teachers who stand firm on their convictions.	1	2	3	4	5	6	7
73.	The completion of any worthwhile task in education requires hard work on the part of pupils.	1	2	3	4	5	6	7
74.	Pupils gain more satisfaction from doing a difficult task well than any other achievement.	1	2	3	4	5	6	7
75.	Pupils never really understand a subject until they can relate what they have learned to the broader problems of the world.	1	2	3	4	5	6	7
76.	Teachers must always be prepared to explain to pupils interrelationships among various elements of the overall curriculum.	1	2	3	4	5	6	7
77.	Teachers must set definite items aside to show pupils the relationships between their subject and the overall goal of education.	1	2	3	4	5	6	7
78.	Pupils gain better understanding of the subject if assignments are presented to them as a series of interrelated problems.	1	2	3	4	5	6	7
79.	The basic function of education is fulfilled only when pupils are led to understand the general significance of the material they have learned.	1	2	3	4	5	6	7
80.	If curriculum plans are to be developed, they must go into detail on how course content can be integrated across subjects.	1	2	3	4	5	6	7
81.	In planning their work teachers should rely heavily on the knowledge and skills pupils have acquired outside the classroom.	1	2	3	4	5	6	7
82.	The pupil's knowledge is best developed when teachers interrelate facts and figures from many different subject fields.	1	2	3	4	5	6	7

	Disagree						Agree
83. The deep interest which pupils sometimes develop in one subject can be valuable to them, but only if teachers succeed in broadening their perspectives across subject matter boundaries.	1	2	3	4	5	6	7
84. The most important thing a teacher can do to set the stage for learning is to discover the interests of students.	1	2	3	4	5	6	7
85. The teacher must avoid strict adherence to the sequence provided by a textbook series.	1	2	3	4	5	6	7
86. An essential component of a good lesson is one of showing how it is related to other areas of knowledge.	1	2	3	4	5	6	7

Appendix F

Instructions

On the next two pages there is a test which consists of pairs of numbers. If the two numbers of a pair are exactly the same, make a check mark (✓) on the line between them; if they are different, make no mark on that line. When the examiner says "Stop", draw a line under the last pair at which you have looked.

Samples done correctly

79542 ____ 79524

5794367 ✓ 5794367

Now try the samples below.

66273894 ____ 66273984

527384578 ____ 527384578

This is a test for Speed and Accuracy. Work as fast as you can without making mistakes.

Do not turn this page until you are told to begin.

√ Check if the two numbers are the same.

1. 307——309
2. 4605——4603
3. 976——996
4. 101267——101267
5. 3065432——30965432
6. 103345700——103345700
7. 46754——466754
8. 3367490——3367490
9. 2779——2778
10. 57394——57394

11. 63801829374——63801839474
12. 283577657——283577657
13. 75689——75689
14. 2547892026——2547893026
15. 336354——336254
16. 998745732——998745733
17. 623——623
18. 263849102983——263849102983
19. 5870——5870
20. 379012——379012

21. 8734629——8734629
22. 2549806746——2549806746
23. 57802564——57892564
24. 689246——688246
25. 1578024683——1578024683
26. 582039485618——582039485618
27. 63829172630——63829172639
28. 592——592
29. 829374820——829374820
30. 62937456——63937456

31. 8293——8293
32. 6382910293——6382910292
33. 781928374012——781928374912
34. 68293——68393
35. 18203649271——18293649271
36. 4820384——4820384
37. 283019283745——283019283745
38. 73927102——73927102
39. 91029354829——91029354829
40. 38291728——38291728

41. 6283910293——6283910203
42. 392018273648——392018273848
43. 820——829
44. 572937273——572937373
45. 7392——7392
46. 8172036——8172036
47. 68391028364——68391028394
48. 48293——48292
49. 739201——739201
50. 62839201——62839211

51. 5829——5820
52. 192836472829——192836472829
53. 362——362
54. 2039271827——2039276837
55. 73829——73829
56. 82739102837——82739102837
57. 48891028——48891028
58. 7291728——7291928
59. 172839102839——172839102839
60. 628192——628102

61. 473829432——473829432
62. 478——478
63. 372816253902——372816252902
64. 64829——64830
65. 4739210249——4739210249
66. 748362——748363
67. 728354792——728354772
68. 3927——3927
69. 927384625——927384625
70. 4628156——4628158

71. 6382——6392
72. 12937453829——12937453829
73. 523——533
74. 7263920——7263920
75. 74293——74293
76. 82734291——82734271
77. 2739102637——2739102637
78. 62810263849——62810263846
79. 638291——638291
80. 62831027——62831027

81. 527——529
82. 172438291026——172438291026
83. 7253829142——725382942
84. 836287——836289
85. 62435162839——62435162839
86. 6254——6256
87. 6241526——6241526
88. 1426389012——1426389102
89. 825——825
90. 67253917287——67253917287

91. 6271——6271
92. 263819253627——263819253629
93. 82637——82937
94. 728392736——728392736
95. 62739——62739
96. 728352689——728352688
97. 463728——463728
98. 73829175——73827196
99. 4825367——4825369
100. 56382018——56382018

√ Check if the two numbers are the same.

101. 789——789
102. 819263728192——819263728172
103. 682537289——682537298
104. 7245——7245
105. 82936542891——82936542891
106. 4738267——4738277
107. 63728——63729
108. 6283628901——6283628991
109. 918264——918264
110. 263728192037——263728192073

111. 52839102738——5283910238
112. 6283——6282
113. 7283529152——7283529152
114. 208——298
115. 82637201927——8263720127
116. 15273826——15273826
117. 72537——73537
118. 726391027384——726391027384
119. 627389——627399
120. 725382910——725382910

121. 46273——46273
122. 629——620
123. 7382517283——7382517283
124. 637281——639281
125. 2738261——2728261
126. 627152637490——627152637490
127. 73526189——73526189
128. 5372——5392
129. 63728142——63728124
130. 4783946——4783046

131. 82637281028——82637281028
132. 628——628
133. 7293728172——7293728177
134. 7362——7362
135. 927382615——927382615
136. 85345——85345
137. 895643278——895642377
138. 726352——726353
139. 7263524——7263524
140. 632685——632685

141. 273648293048——273648293048
142. 634——634
143. 7362536478——7362536478
144. 7362——7363
145. 7362819273——7362819273
146. 63728——63738
147. 63728192637——63728192639
148. 728——738
149. 62738291527——62738291529
150. 63728192——63728192

151. 73526——73526
152. 7283627189——7283627189
153. 627——637
154. 728352617283——728352617282
155. 6281——6381
156. 936271826——936371826
157. 82637192037——82637192037
158. 527182——527182
159. 6273——6273
160. 726354256——72635456

161. 725361552637——725361555637
162. 7526378——7526377
163. 685——685
164. 82637481028——82637481028
165. 3427——3429
166. 827364933251——827364933351
167. 63728——63728
168. 6273846273——6273846293
169. 62836——6283
170. 2638496——2638496

171. 738291627874——738291627874
172. 62826454——62836455
173. 42738267——42738269
174. 573829——573829
175. 628364728——628364928
176. 725——735
177. 627385——627383
178. 63354——63354
179. 54283902——54283602
180. 7283562781——7283562781

181. 62738——63738
182. 727355542321——72735542321
183. 263849332——263849332
184. 162837——163837
185. 47382912——47382922
186. 628367299——628367399
187. 111——111
188. 11829304829——11828304829
189. 4448——4448
190. 333693678——333693678

191. 3212——3212
192. 27389223678——27389223678
193. 473829——473829
194. 7382937——7383937
195. 3628901223——3628901233
196. 5572867——5572867
197. 87263543——87263543
198. 3678902——3678892
199. 15672839——15672839
200. 9927382——9927382

Coder Number_____

Teacher Number_____

Appendix G

BACS-9-Spring 1978

The following statements refer to your opinions regarding the teacher and class you have just seen. After completing the OScAR ratings on this teacher, please indicate your opinion about each of the following statements by circling a number from 1 to 7 to the right of the statement. If you strongly disagree, circle the low number. If you strongly agree, circle the high number. The numbers at the two ends of the scale represent the stronger feelings. Mark a 1 or 7 only when your feelings are exceptionally strong. Please give an opinion about every statement.

	Disagree						Agree
1. The students responded very well to the teacher.	1	2	3	4	5	6	7
2. This was a very effective teacher.	1	2	3	4	5	6	7
3. The teacher had a well prepared lesson plan.	1	2	3	4	5	6	7
4. The teacher liked his/her students.	1	2	3	4	5	6	7
5. The students learned a lot from the teacher.	1	2	3	4	5	6	7
6. This teacher's style of teaching appealed to me.	1	2	3	4	5	6	7